Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 6201–6205, August 1989
Evolution

Date of the monocot–dicot divergence estimated from chloroplast DNA sequence data

(age of angiosperms/phylogeny of cereals/molecular clock)

KENNETH H. WOLFE*†, MANOLO GOUY*‡, YAU-WEN YANG§, PAUL M. SHARP†, AND WEN-HSIUNG LI*¶

*Center for Demographic and Population Genetics, University of Texas, P.O. Box 20334, Houston, TX 77225; †Department of Genetics, Trinity College, Dublin 2, Ireland; ‡Laboratoire de Biométrie, Université Lyon I, 69622 Villeurbanne Cedex, France; and §Department of Cell Biology, Baylor College of Medicine, Houston, TX 77030

Communicated by Robert K. Selander, May 8, 1989 (received for review November 21, 1988)

ABSTRACT **The divergence between monocots and dicots represents a major event in higher plant evolution, yet the date of its occurrence remains unknown because of the scarcity of relevant fossils. We have estimated this date by reconstructing phylogenetic trees from chloroplast DNA sequences, using two independent approaches: the rate of synonymous nucleotide substitution was calibrated from the divergence of maize, wheat, and rice, whereas the rate of nonsynonymous substitution was calibrated from the divergence of angiosperms and bryophytes. Both methods lead to an estimate of the monocot–dicot divergence at 200 million years (Myr) ago (with an uncertainty of about 40 Myr). This estimate is also supported by analyses of the nuclear genes encoding large and small subunit ribosomal RNAs. These results imply that the angiosperm lineage emerged in Jurassic–Triassic time, which considerably predates its appearance in the fossil record (≈120 Myr ago). We estimate the divergence between cycads and angiosperms to be ≈340 Myr, which can be taken as an upper bound for the age of angiosperms.**

The fossil record shows a vast increase in the numbers and distribution of angiosperm species in the mid-Cretaceous period, around 100 million years (Myr) ago (1). The earliest reliable angiosperm macrofossils are about 120 Myr old, but because these are already clearly divisible into monocotyledonous and dicotyledonous types it seems that the earliest stages of angiosperm evolution evaded fossilization (2, 3). Although it is generally accepted that angiosperms descended from the progymnosperm lineage, there is little agreement as to when they arose or even from which branch of the gymnosperms they stem (4). Since the progymnosperm lineage extends back to at least 370 Myr ago (2), there is an enormous range of time during which angiosperms might have had their beginnings. Theories as to why there are no fossils of progenitor angiosperms fall into two basic types: either angiosperms did not exist until the early Cretaceous and then radiated explosively, or pre-Cretaceous angiosperms lived in habitats so refractory to fossilization that they left no record (3–5). In this paper we attempt to decide between these alternative theories by analyzing plant DNA sequences, which can be used to estimate the date of divergence of monocots and dicots, and hence to provide a minimal age for angiosperms themselves.

Despite promising early results from protein sequencing (for example, see ref. 6), molecular data have not been used extensively to investigate plant evolution. An initial application of DNA sequences to studying the origin of angiosperms has recently been made by Martin *et al.* (7). Their analysis is based on comparison between plants, animals, and fungi of the sequences of the nuclear gene (called *gapC* in plants) encoding cytosolic glyceraldehyde-3-phosphate dehydrogenase (GAPDH). By using several divergence dates between animal taxa and between the plant, animal, and fungal kingdoms they were able to calculate the rate of evolution of this gene. From this, they estimated that monocots and dicots diverged about 320 Myr ago, which is long before angiosperms appear in the fossil record. In fact, the earliest land plant fossils are only about 420 Myr old (8, 9), so their result would imply that the lineages leading to bryophytes, pteridophytes, gymnosperms, monocots, and dicots all appeared within the first 100 Myr of land plant evolution. However, their result must be treated with caution because it is based on data from only one gene. This has prompted us to investigate whether the sequences of other genes also point to such an ancient date for the monocot–dicot divergence.

Our analysis is based on the comparison of chloroplast DNA (cpDNA) sequences. These are more useful than nuclear sequences because of their lower rate of silent nucleotide substitution and because far more data are available. The date of divergence of monocots and dicots can be calculated by extrapolation from the known dates of other speciation events by means of phylogenetic reconstruction based on the evolutionary distances between DNA sequences. To calibrate our phylogenetic trees, we have used events both subsequent to and prior to the monocot–dicot split: the divergence of the monocots maize and wheat (approximately 60 Myr ago), and the divergence between bryophytes and the ancestors of angiosperms (approximately 400 Myr ago). From these data we calculate that the divergence between monocots and dicots occurred around 200 Myr ago, which supports the theory that angiosperms existed for a considerable period before coming to dominate the Earth's flora. Comparison of sequences from the nuclear large subunit (26S) and small subunit (18S) ribosomal RNA (rRNA) genes leads to similar estimates of the monocot–dicot divergence date, which suggests that the GAPDH results cannot be generalized to other nuclear genes.

DATA AND METHODS

References to the original publications of the DNA sequences used are given in recent compilations (10–12). Differences between pairs of protein-coding genes were calculated in terms of the numbers of substitutions per site at synonymous positions (K_S) and at nonsynonymous positions (K_A), using the method of Li *et al.* (13). The divergence (K) between rRNA genes was calculated by Kimura's two-parameter method (14). Unrooted phylogenetic trees were then drawn

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Myr, million years; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; cpDNA, chloroplast DNA; K_S and K_A, numbers of substitutions per site at synonymous positions and at nonsynonymous positions, respectively; K, divergence.

¶To whom reprint requests should be addressed.

from these distance measures by the neighbor-joining method (15), which allows for unequal rates of evolution in different lineages. In cpDNA comparisons we excluded genes that are located in the inverted repeat region in some species, but are single-copy in others, because this region appears to have experienced a lower substitution rate than the rest of the chloroplast genome (16). For simplicity, we refer to all nuclear rRNA genes as 26S (large subunit) or 18S (small subunit; genes) these may not be their exact sedimentation coefficients in some species (11, 12).

RESULTS

We present results from three analyses in which the divergence between monocots and dicots can be compared to the divergence between other taxa—namely, (*i*) maize and wheat; (*ii*) angiosperms and bryophytes; and (*iii*) plants, animals, and fungi. In the first case we compare synonymous nucleotide substitutions in chloroplast genes because the distances between the species are small enough that saturation has not occurred. The angiosperm vs. bryophyte comparisons utilize nonsynonymous substitutions in chloroplast protein-coding genes and substitutions in chloroplast-encoded rRNA genes. The third comparison is of nuclear rRNA sequences from plants, animals, and fungi and is intended to investigate whether chloroplast and nuclear genes lead to consistent estimates of the monocot–dicot divergence date.

Calibration Based on the Maize–Wheat Divergence. The use of the divergence between maize and wheat as a reference point for the monocot–dicot divergence date presents a problem in that although a reasonable number of cpDNA sequences can be compared between maize and wheat, the fossil evidence for their date of divergence is rather poor. We overcome this problem by first demonstrating (from molecular data) that maize, wheat, and rice all originated at approximately the same time. This allows us to use the fossil record of rice to estimate the date of this event and hence to date the monocot–dicot split.

Fig. 1 is a phylogenetic tree for maize, wheat, barley, rice, and three dicot species, based on the numbers of synonymous substitutions in three chloroplast genes that have been sequenced in all seven species. Traditionally, wheat and barley are classified into the subfamily Pooideae, whereas maize and rice are in the Panicoideae and Oryzoideae, respectively (17). Opinions differ concerning the relationships among these three subfamilies: Watson *et al.* (17), using a numerical taxonomy approach, propose maize as an outgroup to wheat and rice, whereas an analysis of geographical distributions by Clayton (18) puts rice as the outgroup. The cpDNA data (Fig. 1) support neither of these and suggest that

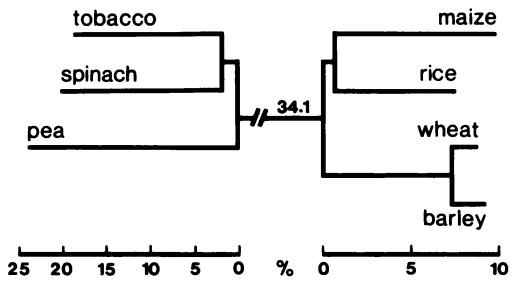

FIG. 1. Phylogenetic tree for four monocot and three dicot species. The tree was inferred by the neighbor-joining method (15) from K_S data for three chloroplast genes: *rbcL*, *atpB*, and *atpE*, totaling 720 synonymous sites. Note that the length (0.7%) of the internal branch leading to the maize–rice pair is less than the standard error (≈1.7%) of the pairwise K_S values between the monocot subfamilies, and thus that branching order is not certain.

the wheat/barley subfamily is the outgroup, but only by a very small margin. In fact, if tobacco alone is used as the reference species, the topology of the tree changes and rice becomes the outgroup. We conclude that the branching order is too close to a trichotomy to be resolved by the available data. This is supported by the nuclear rRNA sequence analysis by Hamby and Zimmer (19), who found that the relative positions of these species within a phylogenetic tree of the grasses changed depending on what method was used to infer the tree. Since fossils of rice leaf epidermis have been described (20) from the upper Eocene (about 40 Myr ago), we take 50 Myr as a lower bound for the origin of maize, wheat, and rice. G. L. Stebbins (cited in refs. 21 and 22) has suggested 65–70 Myr ago as an upper bound for the maize–wheat divergence. Fig. 1 also confirms that wheat and barley are closely related, with a divergence date approximately one-fifth of the maize–wheat–rice divergence, or 10–14 Myr ago.

Table 1 lists the pairwise numbers of synonymous substitutions per site (K_S) between maize, wheat (or its close relatives barley and rye), and tobacco, for 12 chloroplast genes. Ideally, we should use an outgroup species in this comparison to see whether the rates of nucleotide substitution have been equal in the monocot and dicot lineages. However, no suitable outgroup is presently available [liverwort cpDNA (24) is too distantly related for synonymous substitutions to be counted accurately], so we must assume that the rates have been equal in the two lineages. From the weighted means in Table 1, the ratio between the dates of the monocot–dicot divergence and the maize–wheat divergence is calculated as $[(0.578 + 0.576)/2]/0.173 = 3.34$. Taking our estimate of the latter date as 50–70 Myr sets the monocot–dicot divergence at 170–230 Myr ago. If our assumption of equal substitution rates is valid, the date of 170 Myr can be taken as a lower bound because it is based on fossil evidence for the emergence of the rice subfamily by 40 Myr ago (20).

Bryophyte cpDNA as an Outgroup. Our second approach to dating the monocot–dicot divergence is to use as a calibration point the divergence between bryophytes and higher plants. Unequivocal fossil data exist for distinct liverworts by 350 Myr ago and for progymnosperms (the presumed ancestors of

Table 1. Numbers of synonymous substitutions per 100 sites (K_S, %) between maize, wheat (or its siblings), and tobacco for 12 chloroplast genes

Gene	Maize vs. wheat	Maize vs. tobacco	Wheat vs. tobacco	L_S*
atpA	14.6 ± 2.2	53.5 ± 5.6	52.3 ± 5.4	344
atpB	17.5 ± 2.5	65.4 ± 6.4	65.3 ± 6.3	347
atpE	19.3 ± 5.2	50.8 ± 10.1	67.3 ± 12.8	88
atpH	8.4 ± 3.9	38.6 ± 9.7	40.9 ± 10.0	64
orf43	7.2 ± 5.2	43.3 ± 15.8	44.4 ± 16.8	29
orf62	14.5 ± 6.2	46.4 ± 13.3	38.7 ± 11.6	45
psaC	29.6 ± 9.0	79.5 ± 18.8	94.8 ± 23.7	52
psbB	14.6 ± 2.2	56.3 ± 5.6	59.1 ± 6.2	343
psbC	18.9 ± 4.1	52.8 ± 8.5	47.8 ± 7.4	153
psbD	20.0 ± 3.2	48.4 ± 5.9	47.5 ± 5.9	239
psbH	20.5 ± 7.4	60.3 ± 14.7	66.9 ± 16.0	51
rbcL	19.9 ± 2.8	70.2 ± 6.9	62.7 ± 6.2	321
Total	17.3 ± 1.0	57.8 ± 2.4	57.6 ± 2.4	2077

The "total" K_S value for each species pair is a mean over all genes, weighted by the number of sites in each gene; its standard error was calculated as described (16). Wheat sequences were not available for some genes, so we used barley *psbC* and *psbD* and rye *psbB* and *orf43*. The barley *psbC* sequence is not full-length (10). We did not include *orf35* because the maize and rye sequences differ only by an insertion/deletion event. The names *orf35* and *orf43* refer to the two short genes located between *psbB* and *psbH*; *orf62* is downstream of *psbDC* (10, 23).

*Number of synonymous sites compared.

angiosperms and gymnosperms) at about 370 Myr ago (2, 3, 9). Seeds have also been described that are about 350 Myr old (25). These ancient dates are close to the ages of the oldest known land plant fossils, which are found in the mid-Silurian (about 420 Myr ago) (8, 9). Some of these fossils may actually represent the common ancestor of both tracheophytes (vascular plants) and bryophytes (9, 26). It seems unlikely that the tracheophyte–bryophyte split occurred before these plants emerged onto dry land, so we can be confident that it falls between 350 and 450 Myr ago. Even if bryophytes descended from pteridophytes as suggested by the analyses of 5S rRNA sequences by Hori *et al.* (27), the divergence between the angiosperm and bryophyte lineages must still have occurred between these dates.

Synonymous codon positions compared between angiosperm and liverwort cpDNA sequences are almost saturated with substitutions (because these taxa separated so long ago), which makes reliable estimation of K_S impossible. Instead, we use the numbers of substitutions at nonsynonymous sites (K_A) in these comparisons. K_A is considerably lower than K_S due to the effect of selective constraints on amino acid sequences, but this also causes K_A to vary considerably among genes. However, data from different genes can still be pooled to draw phylogenetic trees. In the angiosperm vs. bryophyte comparisons we can also make use of outgroup data (from the green alga *Chlamydomonas* or the protist *Euglena*) to investigate whether the rates of evolution have been equal in different branches of the phylogenetic tree. Whenever possible, we use *Chlamydomonas* as an outgroup in preference to *Euglena* because its cpDNA is more closely related to that of land plants.

Table 2 shows the results of K_A calculations between a monocot (usually maize), tobacco, and liverwort, with *Chlamydomonas* as the outgroup for nine genes and *Euglena* as the outgroup for a further seven genes. Although K_A varies by >10-fold among the genes studied, the relative values of K_A between each species pair are reasonably consistent. Table 2 also shows the results for two chloroplast rRNA genes compared between the same species, with the green algae *Chlamydomonas* and *Chlorella* as outgroups. The unrooted phylogenetic trees in Fig. 2 *a* and *b* were constructed from the weighted mean K_A values in Table 2 using the neighbor-joining method (15). These trees and the trees drawn from the rRNA sequences (Fig. 2 *c* and *d*) show a slower rate of substitution in the tobacco lineage than in the monocot lineage; the average difference in branch lengths is about 2-fold. By using a similar data set, we have shown that the rate of nonsynonymous substitution in tobacco and spinach is significantly slower than in another dicot, pea (16). From Fig. 2 it is also clear that the rate of evolution in the liverwort lineage has been slower than in the angiosperms; the branch length difference is between 1.2- and 2.8-fold, depending on which gene and lineage are considered. These vagaries of the molecular clock make the estimation of divergence times difficult. For example, from the tree in Fig. 2*a*, the ratio between the dates of the monocot–dicot and angiosperm–bryophyte divergences can be calculated as either 0.55 (using the monocot branch) or 0.42 (using the tobacco branch). It is impossible to say which of these estimates is the more reliable because we cannot tell whether the rate difference is due to an acceleration in the monocot lineage or to a slowdown in tobacco. We previously favored the latter hypothesis (16), but the slow rate now also found in liverwort makes an acceleration in monocots (and pea) equally likely.

The branch length ratios from Fig. 2*a* lead to an estimate of 150–250 Myr for the divergence of monocots and dicots, based on a date of 350–450 Myr for the angiosperm–bryophyte divergence. Similarly, the branch lengths of the trees in Fig. 2 *b–d* yield estimates of 160–230, 140–280, and 170–300 Myr, respectively. These are quite consistent with one another and are in good agreement with the result of the maize/wheat comparison. Combining the data in Fig. 2 gives weighted mean values of 0.57 and 0.42 for the ratio calculated between the dates of the monocot–dicot and angiosperm–bryophyte divergences, using the monocot and tobacco lineages, respectively. This leads to an overall estimate of 150–260 Myr based on calibration from the angiosperm–bryophyte divergence. This ratio (about 0.5) is independent of calibration from the fossil record, so if Martin *et al.*'s estimate (7) of about 320 Myr for the monocot–dicot divergence date is correct it implies that the angiosperm and bryophyte lineages separated some 600 Myr ago. This would imply that the two lineages had colonized dry land independently.

Table 2. Divergences between monocots (Mon), tobacco (Tob), liverwort (Liv), and an outgroup species (Out) for 16 protein-coding genes and 2 rRNA genes in cpDNA

Gene	Mon/ Tob	Mon/ Liv	Tob/ Liv	Mon/ Out	Tob/ Out	Liv/ Out	L^*
			Chlamydomonas as outgroup				
atpA	10.4	16.8	7.7	28.3	22.2	19.6	258
atpB	4.5	7.4	7.0	12.6	13.7	12.7	1126
atpE	16.1	32.2	20.6	45.3	46.7	45.2	312
psaA	1.8	4.4	3.7	9.2	8.8	6.6	1749
psaB	1.5	4.3	3.8	11.2	10.3	9.4	1719
psbA	1.0	1.1	1.6	6.9	6.6	7.4	821
psbD	1.0	1.7	1.6	4.2	4.2	3.7	814
rbcL	5.4	6.2	5.1	8.7	8.0	6.4	1095
rpl16	9.3	15.5	11.9	22.7	20.5	16.1	307
Total	3.5	6.3	5.0	11.8	11.3	10.0	8200
			Euglena as outgroup				
atpH	0.6	1.1	0.6	22.4	22.6	22.2	173
orf38	1.1	3.4	4.7	10.1	11.5	8.6	90
psbC	0.7	1.7	1.1	11.3	12.0	11.6	743
psbE	1.0	5.4	5.4	16.1	16.8	13.1	194
psbF	1.1	2.3	1.1	9.2	8.0	8.9	90
rps3	16.0	28.5	23.9	57.3	56.8	52.4	503
rps4	13.3	21.8	22.5	35.1	34.4	32.4	65
Total	5.3	10.1	8.6	26.0	26.2	24.3	1858
			Chlorella as outgroup				
23S rRNA	4.0	6.8	5.2	18.8	16.9	16.4	2666
			Chlamydomonas as outgroup				
16S rRNA	3.4	5.2	3.2	17.7	17.2	15.8	1462

For protein-coding genes, the divergences shown are numbers of substitutions per 100 nonsynonymous sites (K_A, %); for rRNA genes, the corrected number of substitutions per 100 sites (Kimura's K) is shown. The *atpA*, *psbC*, and *rps4* comparisons are for partial sequences. The monocot species is maize for all but four genes (barley *psbAEF*, *orf38*). *orf38* (also known as *psbI* or *psbL*) is a gene immediately downstream of *psbEF* (10, 23). We have not included *rpl22* in the comparisons because of its very high rate of evolution (K_A between maize and tobacco is 40%); there are also large insertions/deletions between species.

*Number of sites compared; for protein-coding genes, $L = L_A$.

Analysis of Nuclear rRNA Genes. Because it is possible that the differences between the dates estimated in our analysis and in that of Martin *et al.* (7) are due to fundamental differences between the evolution of nuclear and chloroplast genomes, we also analyzed the sequences of two nuclear genes: the 26S and 18S rRNAs. Phylogenetic trees for these genes with animals and fungi as outgroups are drawn in Fig. 3. We have recently proposed from analyses of several macromolecular sequences that fungi are an outgroup to plants and animals by a small margin (28). If we assume, rather arbitrarily as did Martin *et al.* (7), that the plant–animal divergence occurred 1000 Myr ago, the branch lengths in Fig. 3 give estimates of 200–250 Myr from the 26S gene and 200–210 Myr from the 18S gene (calculated from the length

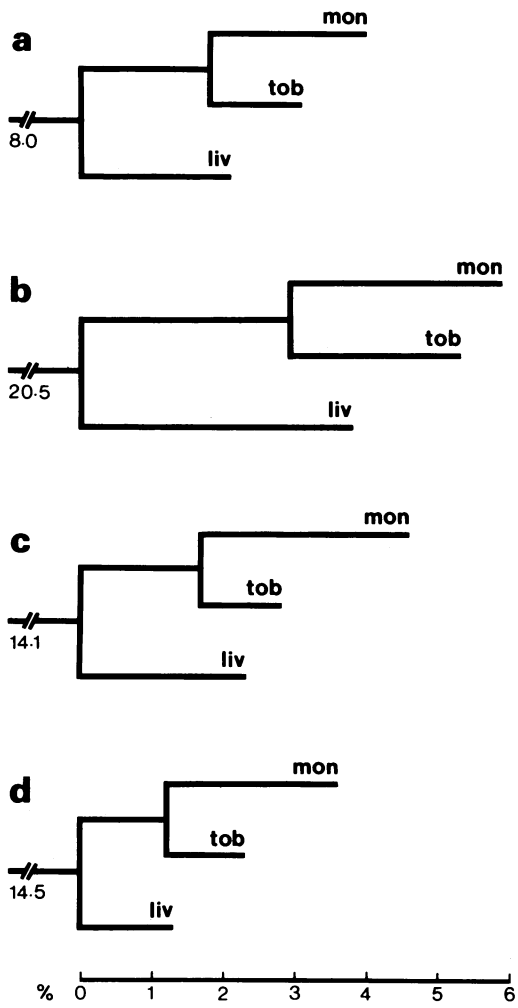

FIG. 2. Phylogenetic trees for monocots (mon), tobacco (tob), and liverwort (liv) based on cpDNA sequences, using green algal or protist species as outgroups. (*a*) Nine protein-coding genes with *Chlamydomonas* as outgroup. (*b*) Seven protein-coding genes with *Euglena* as outgroup. (*c*) The 23S rRNA with *Chlorella* as outgroup. (*d*) The 16S rRNA with *Chlamydomonas* as outgroup. The trees are drawn from the data in Table 2.

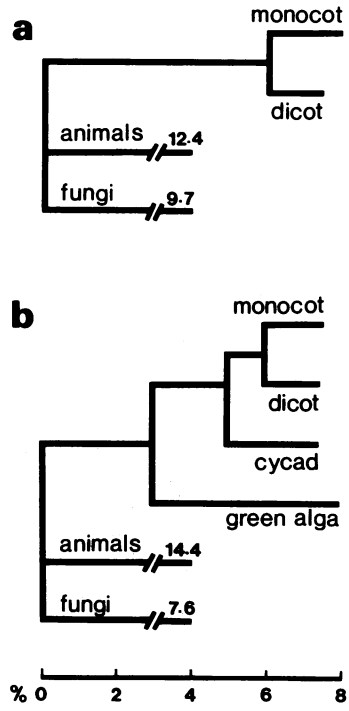

FIG. 3. Phylogenetic trees for plant, animal, and fungal nuclear 26S (*a*) and 18S (*b*) rRNA sequences. Only conserved regions of the molecules were compared, totaling 2323 base pairs for the 26S gene and 1564 base pairs for the 18S gene. The 26S sequences are from rice, lemon, human, *Xenopus*, *Caenorhabditis*, and *Saccharomyces*. The 18S sequences are from rice, maize, soybean, *Zamia* (cycad), *Chlamydomonas*, *Nanochlorum*, human, mouse, rabbit, *Xenopus*, *Artemia*, *Caenorhabditis*, *Saccharomyces*, and *Neurospora*. Sequence data can be found in GenBank (release 58) and refs. 11 and 12.

of the monocot or dicot branch, in relation to the total length back to the plant–animal–fungal branch point). These agree with the estimates from cpDNA data.

From the 18S rRNA tree (Fig. 3*b*) we can also infer branching dates for cycads and green algae. We estimate the origin of cycads to be 340 Myr ago, which predates their fossil appearance by about 100 Myr (2). Although a controversial issue, the 5S rRNA data of Hori *et al.* (27) suggest that gymnosperms (as represented by cycads, conifers, and *Ginkgo*) are monophyletic, in which case the cycad branching date applies to other gymnosperms. Fig. 3*b* suggests a date of 630 Myr for the divergence between land plants and green algae (*Chlamydomonas* and *Nanochlorum*). This seems reasonable in that the green algal lineage (from which land plants were derived) extends back into the Precambrian (2, 29), but there is no direct fossil record of the green alga–land plant divergence with which our estimate can be compared.

DISCUSSION

When using the divergence between maize and wheat to calculate the monocot–dicot divergence date we assumed an equal rate of synonymous substitution in the monocot and tobacco cpDNA lineages. This assumption needs to be substantiated because it is clear that there is a difference in the rate of substitution in monocots and dicots at nonsynonymous sites and in chloroplast rRNA genes. Such a concerted change in nonneutral substitution rates could indicate a change in mutation rate, in which case the synonymous substitution rate should also be affected. However, in a previous analysis of similar data (16), we found that the slowdown in tobacco (with respect to pea) was much less pronounced in its effects on K_S than on K_A. This suggests that the synonymous rates in monocots and tobacco may not be too unequal. Furthermore, comparisons with a gymnosperm *rbcL* sequence (from Douglas-fir; S. H. Strauss, personal communication) show equal synonymous substitution rates in monocots and dicots for this gene. The rate change in K_A but not K_S is puzzling: if the underlying mutation rate is constant, a given protein is expected to evolve at a fixed rate, determined by the constraints on its amino acid sequence. One possible explanation is that negative selection on amino acid replacements is more effective in tobacco, which could occur if it has a larger effective population size (30). However, no rate differences are seen when tobacco and pea mitochondrial sequences are compared to monocot sequences (not shown). It might be preferable to use pea as a representative dicot in place of tobacco, but this would entail a reduction in the amount of data analyzed and would lead to further possible complications arising from the lack of an inverted repeat in pea cpDNA. In the absence of useful outgroups (such as gymnosperm sequences for all genes in Table 1), it is necessary to assume equal rates of synonymous substitution.

Combining the estimates from the two sets of cpDNA data, we propose that monocots and dicots diverged 200–205 Myr ago (with an uncertainty of about 40 Myr), probably in the late

Triassic. Although many angiosperm subfamilies were distinguishable 90 Myr ago, there are almost no reliable angiosperm macrofossils older than 120 Myr (1). There is questionable fossil evidence for late Triassic monocots and for Jurassic dicots (2, 4). The date of the monocot–dicot divergence is necessarily an underestimate of the age of angiosperms themselves. Thus our result points to an origin of angiosperms in the Triassic or even earlier and supports the hypothesis that angiosperms existed long before they came to prominence (3–5). Our estimated date for the divergence between cycads and angiosperms (340 Myr; early Carboniferous) can be taken as an upper bound, so that the origin of angiosperms can be inferred to be between 200 and 340 Myr ago.

Although our result confirms Martin *et al.*'s conclusion (7) that DNA sequences indicate a pre-Cretaceous origin of angiosperms, our estimated date for the monocot–dicot divergence is over 100 Myr more recent than theirs. Our estimate is likely to be more reliable because it is based on over 16,000 nucleotide sites (Tables 1 and 2) as against about 780. However, it is also possible that the evolutionary histories of nuclear DNAs and cpDNAs have been different during plant evolution, which could lead to different estimates from different data sets. To investigate this, we have also estimated the date of the monocot–dicot divergence by using the nuclear rRNA genes (Fig. 3). These genes suggest a date of 200–250 Myr, which is in good agreement with our estimate from cpDNA sequences, but considerably less than the estimate of 320 Myr obtained from the nuclear GAPDH sequences (7). From this it appears that the latter estimate is peculiar to GAPDH, and there is probably no real disagreement between molecular phylogenies derived from nuclear DNAs and cpDNAs.

There is no *a priori* reason why either nuclear or organelle DNA should be more useful in evolutionary studies of this kind. We have found chloroplast sequences to be preferable for several reasons: far more data are presently available, more closely related outgroup species can be used, and the slow rate of evolution of cpDNA allows synonymous and nonsynonymous sites to be compared between monocots and dicots. The major problem encountered with cpDNA sequences is unequal rates of evolution in different lineages. At present, we do not know whether this is also true of plant nuclear or mitochondrial genes, although there seems to be a good molecular clock for nuclear 18S rRNA sequences in plants (Fig. 3*b*). The availability of cyanobacterial DNA sequences as references may allow even deeper branchings in the evolutionary history of plants to be investigated using cpDNA data (e.g., ref. 31). It is possible that plant mitochondrial DNA, with its extremely low rate of point mutation (16), could provide even more accurate estimates of the dates of major events in plant evolution. However, the radically different structure of mitochondrial DNA in green algae (32) may mean that there are considerable rate differences between lineages in such comparisons.

We are grateful to Bill Martin and Quentin Cronk for helpful comments and to Steve Strauss for making unpublished data available. This study was supported by National Institutes of Health Grant GM30998.

1. Lidgard, S. & Crane, P. R. (1988) *Nature (London)* **331,** 344–346.
2. Stewart, W. N. (1983) *Paleobotany and the Evolution of Plants* (Cambridge Univ. Press, Cambridge, U.K.).
3. Thomas, B. A. & Spicer, R. A. (1987) *The Evolution and Palaeobiology of Land Plants* (Croom Helm, London).
4. Doyle, J. A. & Donoghue, M. J. (1986) *Bot. Rev.* **52,** 321–431.
5. Friis, E. M., Chaloner, W. G. & Crane, P. R. (1987) in *The Origins of Angiosperms and Their Biological Consequences*, eds. Friis, E. M., Chaloner, W. G. & Crane, P. R. (Cambridge Univ. Press, Cambridge, U.K.), pp. 1–15.
6. Ramshaw, J. A. M., Richardson, D. L., Meatyard, B. T., Brown, R. H., Richardson, M., Thompson, E. W. & Boulter, D. (1972) *New Phytol.* **71,** 773–779.
7. Martin, W., Gierl, A. & Saedler, H. (1989) *Nature (London)* **339,** 46–48.
8. Edwards, D. & Feehan, J. (1980) *Nature (London)* **287,** 41–42.
9. Gensel, P. G. & Andrews, H. N. (1984) *Plant Life in the Devonian* (Praeger, New York).
10. Wolfe, K. H. (1989) *Plant Mol. Biol. Rep.* **7,** 30–48.
11. Dams, E., Hendriks, L., Van de Peer, Y., Neefs, J.-M., Smits, G., Vandenbempt, I. & De Wachter, R. (1988) *Nucleic Acids Res.* **16,** r87–r173.
12. Gutell, R. R. & Fox, G. E. (1988) *Nucleic Acids Res.* **16,** r175–r269.
13. Li, W.-H., Wu, C.-I. & Luo, C.-C. (1985) *Mol. Biol. Evol.* **2,** 150–174.
14. Kimura, M. (1980) *J. Mol. Evol.* **16,** 111–120.
15. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425.
16. Wolfe, K. H., Li, W.-H. & Sharp, P. M. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 9054–9058.
17. Watson, L., Clifford, H. T. & Dallwitz, M. J. (1985) *Aust. J. Bot.* **33,** 433–484.
18. Clayton, W. D. (1981) *Ann. Mo. Bot. Gard.* **68,** 5–14.
19. Hamby, R. K. & Zimmer, E. A. (1988) *Plant Syst. Evol.* **160,** 29–37.
20. Stebbins, G. L. (1981) *Ann. Mo. Bot. Gard.* **68,** 75–86.
21. Zurawski, G., Clegg, M. T. & Brown, A. H. D. (1984) *Genetics* **106,** 735–749.
22. Chao, S., Sederoff, R. & Levings, C. S., III (1984) *Nucleic Acids Res.* **12,** 6629–6644.
23. Wolfe, K. H. & Sharp, P. M. (1988) *Gene* **66,** 215–222.
24. Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. & Ozeki, H. (1986) *Nature (London)* **322,** 572–574.
25. Gillespie, W. H., Rothwell, G. W. & Scheckler, S. E. (1981) *Nature (London)* **293,** 462–464.
26. Remy, W. (1982) *Science* **215,** 1625–1627.
27. Hori, H., Lim, B.-L. & Osawa, S. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 820–823.
28. Gouy, M. & Li, W.-H. (1989) *Mol. Biol. Evol.* **6,** 109–122.
29. Schopf, J. W., Hayes, J. M. & Walter, M. R. (1983) in *Earth's Earliest Biosphere: Its Origin and Evolution*, ed. Schopf, J. W. (Princeton Univ. Press, Princeton, NJ), pp. 361–384.
30. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, U.K.).
31. Ritland, K. & Clegg, M. T. (1987) *Am. Nat.* **130,** S74–S100.
32. Gray, M. W. & Boer, P. H. (1988) *Philos. Trans. R. Soc. London Ser. B* **319,** 135–147.